Genes and Immunity (2012) 13, 232–238

www.nature.com/gene

ORIGINAL ARTICLE

Role of *MYH9* and *APOL1* in African and non-African populations with lupus nephritis

CP Lin[1], I Adrianto[1], CJ Lessard[1,2], JA Kelly[1], KM Kaufman[1,2,3], JM Guthridge[1], BI Freedman[4], J-M Anaya[5], ME Alarcón-Riquelme[1,6], on behalf of the BIOLUPUS and GENLES Networks, BA Pons-Estel[7], J Martin[8], S Glenn[1], A Adler[1], S-C Bae[9], S-Y Park[9], S-Y Bang[9], Y-W Song[10], SA Boackle[11], EE Brown[12], JC Edberg[12], GS Alarcón[12], MA Petri[13], LA Criswell[14], R Ramsey-Goldman[15], JD Reveille[16], LM Vila[17], GS Gilkeson[18], DL Kamen[18], J Ziegler[19], CO Jacob[20], A Rasmussen[1], JA James[1,2], RP Kimberly[12], JT Merrill[21], TB Niewold[22], RH Scofield[1,2,3], AM Stevens[23,24], BP Tsao[25], TJ Vyse[26], CD Langefeld[19], KL Moser[1,2], JB Harley[27,28], PM Gaffney[1] and CG Montgomery[1]

Systemic lupus erythematosus (SLE) is a complex autoimmune disease characterized by autoantibody production and organ damage. Lupus nephritis (LN) is one of the most severe manifestations of SLE. Multiple studies reported associations between renal diseases and variants in the non-muscle myosin heavy chain 9 (*MYH9*) and the neighboring apolipoprotein L 1 (*APOL1*) genes. We evaluated 167 variants spanning *MYH9* for association with LN in a multiethnic sample. The two previously identified risk variants in *APOL1* were also tested for association with LN in European-Americans (EAs) ($N = 579$) and African-Americans (AAs) ($N = 407$). Multiple peaks of association exceeding a Bonferroni corrected *P*-value of $P < 2.03 \times 10^{-3}$ were observed between LN and *MYH9* in EAs ($N = 4620$), with the most pronounced association at rs2157257 ($P = 4.7 \times 10^{-4}$, odds ratio (OR) = 1.205). A modest effect with *MYH9* was also detected in Gullah (rs8136069, $P = 0.0019$, OR = 2.304). No association between LN and *MYH9* was found in AAs, Asians, Amerindians or Hispanics. This study provides the first investigation of *MYH9* in LN in non-Africans and of *APOL1* in LN in any population, and presents novel insight into the potential role of *MYH9* in LN in EAs.

Genes and Immunity (2012) **13,** 232–238; doi:10.1038/gene.2011.82; published online 22 December 2011

Keywords: *MYH9*; *APOL1*; lupus nephritis; systemic lupus erythematosus; multiethnic association study

INTRODUCTION

Systemic lupus erythematosus (SLE) is a prototypic systemic autoimmune disease characterized by multisystem involvement and the development of an immune response against self-antigens, leading to tissue inflammation, destruction and often end-organ damage. SLE is more prevalent in females compared with males (9:1) and in African-American (AA), Asian (AS) and Hispanic (HI) populations compared with EuropeanAmericans (EA).[1–3] Patients classified with SLE manifest a minimum of 4 out of 11 criteria set by the American College of Rheumatology[4,5] with neurological, renal and hematological manifestations representing more severe disease. Lupus nephritis (LN) is one of the most severe complications, drastically increasing the morbidity and mortality of SLE patients,[6] with up to 60% of adult and 80% of pediatric SLE cases developing renal abnormalities during the course of the disease.[7,8] The incidence of LN is higher in AA, HI and AS compared with populations of European ancestry: one study showed that incidences of renal disease for AA and HI are 68.9% and 60.6%, respectively, compared with EA (29.1%) after 5.5 years of follow-up;[9] a similar elevated incidence in AS has also been confirmed.[10,11]

Renal dysfunction is not an exclusive manifestation of SLE; it is a feature of a large number of diseases that may share underlying mechanisms or predisposing genetic factors. There have been recent reports of genetic association of variants located within the non-muscle myosin heavy chain 9 (*MYH9*) gene on chromosome 22 and a variety of renal diseases including focal segmental glomerulosclerosis (FSGS), HIV-associated nephropathy, hypertension-attributed end-stage renal disease (H-ESRD) and diabetic and non-diabetic ESRD in African-derived populations.[12–15]

[1]Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA; [2]College of Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA; [3]US Department of Veterans Affairs Medical Center, Oklahoma City, OK, USA; [4]Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA; [5]Center for Autoimmune Diseases Research (CREA), Universidad del Rosario, Bogota, Colombia; [6]Pfizer, Universidad de Granada - Junta de Andalucía Centre for Genomics and Oncological Research (GENYO), Granada, Spain; [7]Sanatorio Parque, Rosario, Argentina; [8]Instituto de Parasitologia y Biomedicina Lopez-Neyra, Consejo Superior de Investigaciones Cientificas (CSIC), Granada, Spain; [9]Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Republic of Korea; [10]Division of Rheumatology, Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea; [11]Division of Rheumatology, University of Colorado Denver, Aurora, CO, USA; [12]Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA; [13]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [14]Rosalind Russell Medical Research Center for Arthritis, University of California San Francisco, San Francisco, CA, USA; [15]Division of Rheumatology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; [16]Rheumatology and Clinical Immunogenetics, University of Texas Health Science Center at Houston, Houston, TX, USA; [17]Department of Medicine, Division of Rheumatology, University of Puerto Rico Medical Sciences Campus, San Juan, Puerto Rico; [18]Department of Medicine, Medical University of South Carolina, Charleston, SC, USA; [19]Department of Biostatistical Sciences, Wake Forest University Health Sciences, Winston-Salem, NC, USA; [20]Department of Medicine, University of Southern California, Los Angeles, CA, USA; [21]Clinical Pharmacology, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA; [22]Section of Rheumatology and Gwen Knapp Center for Lupus and Immunology Research, University of Chicago, Chicago, IL, USA; [23]Division of Rheumatology, Department of Pediatrics, University of Washington, Seattle, WA, USA; [24]Center for Immunity and Immunotherapies, Seattle Children's Research Institute, Seattle, WA, USA; [25]Division of Rheumatology, Department of Medicine, University of California Los Angeles, Los Angeles, CA, USA; [26]Divisions of Genetics and Molecular Medicine and Immunology, Infection and Inflammatory Disease, King's College London, London, UK; [27]Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA and [28]US Department of Veterans Affairs Medical Center Cincinnati, Cincinnati, OH, USA. Correspondence: Dr CG Montgomery, Arthritis and Clinical Immunology Research Program, Oklahoma Medical Research Foundation, 825 N.E. 13th Street, MS #57, Oklahoma City, OK 73104, USA. E-mail: Courtney-Montgomery@omrf.org

Received 9 May 2011; revised 4 November 2011; accepted 14 November 2011; published online 22 December 2011

In addition, several monogenic syndromes with point mutations in *MYH9* have been characterized by thrombocytopenia, leukocyte abnormalities and renal failure.[16] Recent studies of FSGS and H-ESRD in AAs,[17,18] however, suggest that the pronounced association of the *MYH9* E-1 risk haplotype (rs4821480, rs2032487, rs4821481 and rs3752462) is primarily due to strong linkage disequilibrium (LD) with two independent genetic variants (rs73885319 and rs71785313) within the neighboring apolipoprotein L1 (*APOL1*) gene.[17-19] Although one study has assessed association between LN and *MYH9* in AAs (and found none),[20] no such study has been conducted for either *MYH9* or *APOL1* in non-African populations. It was therefore the aim of this study to investigate the role of these genes in African and non-African SLE populations with LN. Specifically, we sought to assess association of LN with *MYH9* and evaluate the association between 2 variants within *APOL1* in EA and AA samples and 167 *MYH9* variants and LN in a large multiethnic group sample comprising of EA, AA, AS, HI, Amerindian and Gullah (a unique AA population from the coastal regions of South Carolina and Georgia) samples.

RESULTS

Association analysis of single-nucleotide polymorphisms (SNPs) within *MYH9* comparing LN cases and healthy controls resulted in no significance in the AA, HI, AS or Amerindian populations (Figure 1, Supplementary Figure 1 and Supplementary Table 5). The EA population yielded multiple SNPs exceeding the Bonferroni correction ($P<2.03\times10^{-3}$) with the significant signals of *P*-value $<10^{-3}$ at rs2157257 ($P=4.7\times10^{-4}$, odds ratio (OR) = 1.205), rs5750250 ($P=5.4\times10^{-4}$, OR = 1.472), rs2413396 ($P=6.74\times10^{-4}$, OR = 1.327) and rs4820232 ($P=9.20\times10^{-4}$, OR = = 1.196), all centered at approximately 35.04 Mb (Figure 1 and Table 1). In addition, two of the E-1 haplotype[14] SNPs (rs4821480, rs2032487) were also below the Bonferroni significance threshold (Table 1). A modest association was also detected in the Gullah (70 LN cases/122 healthy controls) at 35.05 Mb, with the strongest association observed at rs8136069 ($P=1.923\times10^{-3}$, OR = 2.304, 95% confidence interval = 1.360 - 3.904). Results of analyses comparing LN cases with SLE cases without LN were very similar, save a slight increase in the number of significant SNPs (Supplementary Table 7a and b).

To further elucidate the effects observed in EA and to determine if any particular SNP was driving the association observed in the region independently, conditional association analyses were performed. Based on LD ($r^2>0.8$) (Figure 2), we incorporated each of the associated SNPs in the region with *P*-value $<1\times10^{-3}$ (rs2413396, rs2157257, rs5750250 or rs4820232) as covariates, one at a time in our model. The effect of rs2413396 became not significant ($P>1\times10^{-1}$) when conditioning on rs5750250 and the converse showed similar results (Table 2). Similarly, the signal diminished at rs2157257 when conditioning on rs4820232. This implies the effects of rs2413396 and rs5750250 or rs2157257 and rs4820232 are non-independent. Both rs2413396 and rs5750250 remained significant upon conditioned on any of rs2157257 and rs4820232B. Likewise, rs2157257 and rs4820232 were still significant after conditioning on either rs2413396 or rs5750250. Furthermore, conditioning on both rs2157257 and rs5750250, the association signals reduced to baseline ($P>1\times10^{-1}$) (Table 2). Therefore, the two main effects exist in the region tagged by (rs2413396, rs5750250) and (rs2157257, rs4820232) as effect 1 and effect 2, respectively, (Figure 2).

In order to differentiate association signals observed between the AA and Gullah populations, we refined our analysis to include only those subjects with >90% African ancestry in the AA (105 LN cases/312 healthy controls) and Gullah (42 LN cases/69 healthy controls) samples. The association signals in the AA and the Gullah samples enriched for African ancestry were less significant than that of the full set of samples (Supplementary Figure 3), perhaps suggesting the influence of European admixture. While we were unable to examine variants within *APOL1* in the Gullah, only a nominal effect at rs71785313 ($P=0.023$) of *APOL1* was seen in the AA sample (Supplementary Table 6). Note too that the association *P*-values of our two genotyped SNPs within the *MYH9* E-1 risk haplotype,[14] known to tag *APOL1* (Supplementary Figure 2), were also insignificant (rs4821481, $P=0.1759$; rs3752462, $P=0.5104$) in AA. These results support our above conclusion of no association between LN and either *MYH9* or *APOL1* in AA.

The previously reported associations of *MYH9* and *APOL1* with kidney disease were described in rather homogeneous phenotypes with severe renal failure (H-ESRD and FSGS); we speculated that the lack of significant association with *MYH9* and *APOL1* in the AA might be due to a wider severity spectrum of renal dysfunction associated with SLE. To explore this possibility, an analysis of a subset of patients for whom we had information on dialysis or kidney transplant (indicating ESRD) was performed. Fisher's exact tests of variants within *MYH9* produced a nominally significant *P*-value of 0.041 at rs10483194 in the EA (30 LN cases/3,491 healthy controls) and $P=0.044$ at rs739095 in the AA (67 LN cases/1,811 healthy controls). However, we found no significant association within the *APOL1* variants (57 LN cases/202 healthy controls) (Supplementary Table 8). We repeated Fisher's exact tests using non-LN SLE as controls, results were not considerably different, except a marginal *P*-value was found at an *APOL1* variant rs71785313 ($P=0.0418$) (Supplementary Table 8). It should be noted that even with this small sample, given the magnitude of the ORs previously reported for *APOL1*,[17] we had greater than 85% power to detect an effect at a *P*-value of 0.01. Thus it suggests there is no effect of *APOL1* and only a slight potential for an effect of *MYH9* in LN, ESRD patients.

Finally, on the basis of a previous report of germline *MYH9* mutations in a patient with SLE and end-stage renal disease,[21] we performed an exploratory analysis to address the potential role of *MYH9* in a broader class of SLE-related target organ damage. Specifically, two subsets of SLE patients of European descent with renal disease and/or thrombocytopenia as well as with renal disease and/or serositis in addition to healthy controls were evaluated for association with *MYH9*. The significance of our strongest effect at rs2413396 was increased when adding cases with thrombocytopenia ($P=2.46\times10^{-4}$, 1351 cases/3491 healthy controls) compared with renal disease alone ($P=6.74\times10^{-4}$, 1129 LN cases/3491 healthy controls), with loss of significance when adding patients with serositis compared with renal disease alone ($P=1.86\times10^{-2}$, 1984 cases/3491 healthy controls) (Supplementary Table 9). The results of the same subgroup analyses using non-LN SLE cases as 'controls' were similar, an increase in significance by two orders of magnitude at rs10483194 was observed when adding cases with thrombocytopenia to the renal diseased patients (1351 cases/1006 non-LN SLE), however, when adding patients with serositis, significance increased marginally as well (1984 cases and 606 non-LN SLE) (Supplementary Table 9).

DISCUSSION

MYH9 encodes the motor protein MYH class II and isoform A, and is expressed mainly in podocytes, peritubular capillaries and tubules of mature kidney. It is responsible for cell polarity, trafficking and cell architecture. Dysregulation may lead to renal complications and eventual glomerulosclerosis. It may be implicated in SLE via its role in phagocytosis of apoptotic leukocytes.[22] Further, *MYH9* is associated with a variety of diverse syndromes that share leukocyte inclusions, abnormally large platelets, thrombocytopenia and bleeding tendency; many of which also include glomerulopathy with progressive renal failure.[23-25] More recently, *MYH9* was implicated in ESRD and FSGS in populations of African ancestry[13,15,26] with even stronger

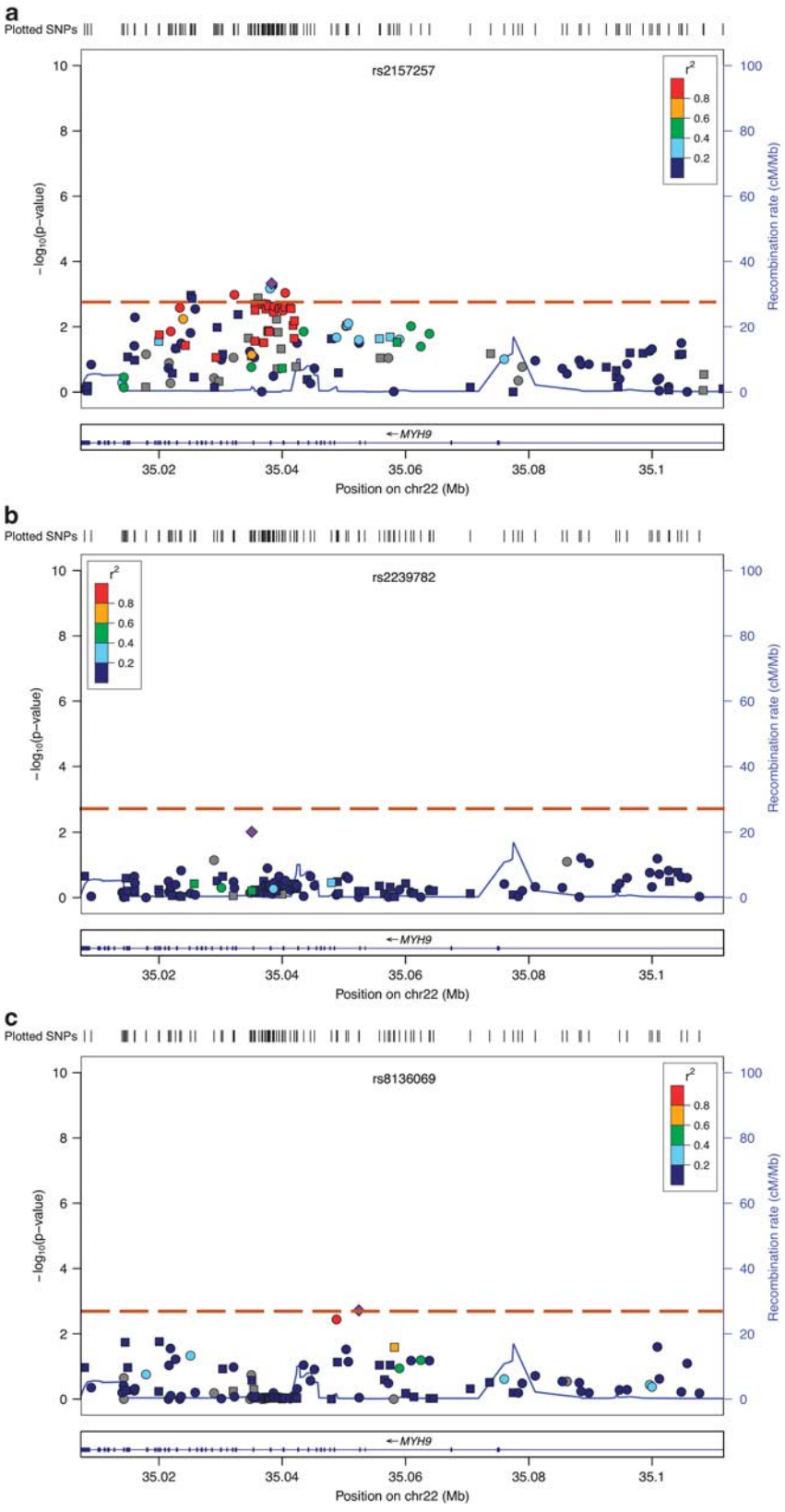

Figure 1. Summary of association analysis for *MYH9* SNPs. Plots of SNPs within the *MYH9* gene associated with lupus nephritis in (**a**) European American, (**b**) African American and (**c**) Gullah populations. Genotyped SNPs are represented by circles and imputed SNPs are shown in rectangles. The recombination rate calculated from the combined CEU, YRI and JPT + CHB Hapmap data is denoted by the purple solid line. The dotted line refers to the Bonferroni threshold of significance.

Table 1. Summary results of the top 20 most significant SNPs from association analyses in European-Americans (LN cases and healthy controls) and their respective odd ratios and confidence intervals for African-Americans and Gullah

SNP[a]	*BP*[b]	*Allele*[c]	*MAF*[d]	*AA OR (95% CI)*[e]	*Gullah OR (95% CI)*[e]	*EA OR (95% CI)*[e]	*EA P-value*[f]
rs9610486	35023388	A/G	0.310	1.011 (0.864,1.183)	1.002 (0.595,1.687)	1.178 (1.059,1.311)	2.627E-03
rs4821480[g]	35025193	T/G	0.043	NA	NA	1.454 (1.161,1.819)	1.085E-03
rs2032487[g]	35025374	T/C	0.043	NA	NA	1.443 (1.153,1.806)	1.364E-03
rs1009150	35032246	A/G	0.296	0.950 (0.818,1.104)	1.046 (0.657,1.665)	1.195 (1.074,1.329)	1.045E-03
rs1557530	35035568	A/G	0.328	1.045 (0.893,1.223)	0.941 (0.539,1.644)	1.177 (1.061,1.306)	2.039E-03
rs5750249	35036062	G/A	0.312	NA	NA	1.191 (1.071,1.325)	1.305E-03
rs28409177	35036217	C/T	0.300	NA	NA	1.180 (1.062,1.312)	2.020E-03
rs2157252	35036825	A/C	0.329	1.028 (0.885,1.194)	1.033 (0.619,1.727)	1.178 (1.061,1.307)	2.096E-03
rs2157254	35037146	C/G	0.327	1.059 (0.912,1.230)	0.995 (0.592,1.671)	1.177 (1.061,1.305)	1.993E-03
rs2157256	35037607	G/A	0.332	1.116 (0.970,1.285)	0.983 (0.612,1.581)	1.177 (1.061,1.305)	2.043E-03
rs9622377	35037819	T/C	0.331	1.056 (0.909,1.227)	1.024 (0.612,1.714)	1.174 (1.059,1.301)	2.321E-03
rs2413396	35038030	T/C	0.090	1.059 (0.924,1.213)	0.974 (0.641,1.481)	1.327 (1.127,1.562)	6.742E-04
rs2157257	35038284	G/A	0.339	NA	NA	1.205 (1.085,1.338)	4.700E-04
rs5750250	35038429	A/G	0.046	1.021 (0.892,1.169)	0.945 (0.620,1.441)	1.471 (1.182,1.830)	5.403E-04
rs4820229	35038699	G/A	0.342	1.074 (0.925,1.247)	1.027 (0.618,1.704)	1.170 (1.060,1.312)	2.392E-03
rs4820230	35039485	A/G	0.327	1.092 (0.947,1.259)	0.961 (0.588,1.571)	1.170 (1.057,1.302)	2.680E-03
rs5756142	35040002	G/A	0.330	1.062 (0.915,1.232)	1.007 (0.602,1.684)	1.171 (1.056,1.299)	2.704E-03
rs4821484	35040465	A/C	0.331	1.063 (0.917,1.233)	0.972 (0.586,1.610)	1.172 (1.057,1.300)	2.519E-03
rs4820232	35040487	G/A	0.341	1.080 (0.934,1.249)	1.020 (0.625,1.663)	1.196 (1.076,1.329)	9.199E-04
rs8141971	35041308	G/A	0.332	1.042 (0.898,1.210)	0.991 (0.594,1.653)	1.172 (1.057,1.300)	2.590E-03

Abbreviations: CI, confidence intervals; EA, European-Americans; LN, lupus nephritis; MAF, minor allele frequency; NA, not applicable; OR, overall survival; SNPs, single-nucleotide polymorphisms. [a]Observed SNPs were in bold. [b]Base pair, location in NCBI36 assembly. [c]Allele (major/minor). [d]MAF. [e]OR was calculated with respect to minor allele and 95% confidence interval for OR. [f]Association results were derived using logistic regression assuming additive mode of inheritance adjusted for global African, Asian and European ancestry estimates and gender. [g]E-1 haplotype SNPs were in italic.

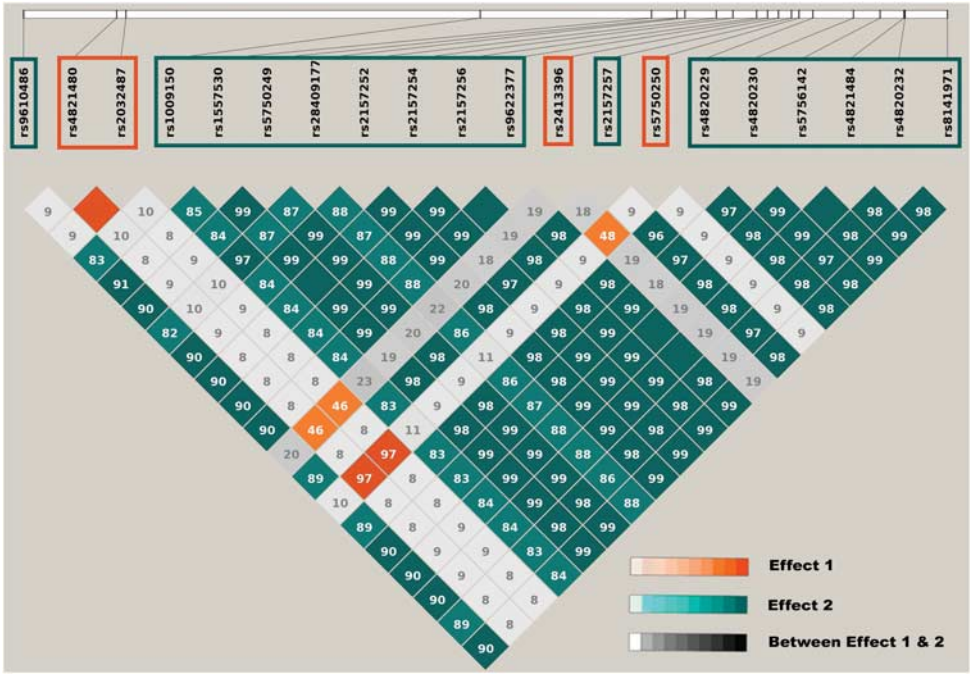

Figure 2. The linkage disequilibrium for significant SNPs in European Americans. This plot shows the pair-wise LD between all markers given in the figure. SNPs in red box represent effect 1, and SNPs in green box represent effect 2. The intensities of the LD between SNPs are depicted in orange (effect 1), green (effect 2) and black (between effect 1 and effect 2).

association with two coding variants in the neighboring *APOL1* gene.[17,18] Its involvement in LN in AA populations has previously been refuted, and the results of this study support this conclusion. However, the role of *MYH9* in LN in non-African populations or of *APOL1* in LN in any ethnic group has not been previously studied.

While potential weaknesses of this report include limited sample size for the Gullah, HI and Amerindian populations and further replication studies are needed, we interrogated the largest number of individuals of these ethnic groups available at the time. Further, the absence of information about dialysis and/or kidney transplant in the medical records for the AS, Gullah, HI and Amerindian limits the conclusions we can draw about those populations.

In summary, this report provides the first evidence of association between LN and *MYH9* variants in a large study

Table 2. Summary results of conditional analysis of the significant SNPs in EA

SNP	*BP*	*Single SNP* P	P *conditional on*				
			rs2413396	*rs2157257*	*rs5750250*	*rs4820232*	*rs2157257 + rs5750250*
rs2413396	35040002	6.74E-04	NA	2.11E-02	1.42E-01	3.11E-02	1.93E-01
rs2157257	35040129	4.70E-04	3.30E-02	NA	8.70E-03	8.01E-01	NA
rs5750250	35040465	5.40E-04	1.61E-01	3.66E-02	NA	1.10E-02	NA
rs4820232	35040487	9.20E-04	3.81E-02	6.40E-01	1.70E-02	NA	6.44E-01

Table 3. Summary of samples passing quality control

Population	*No. of samples*	*Healthy controls*	*Lupus nephritis cases*	*Non-lupus nephritis SLE cases*
African-American	3104	1811	634	659
Asian	2399	1260	529	610
European-American	6781	3491	1129	2161
Gullah	272	123	70	79
Hispanic	1274	336	439	499
Amerindian	937	471	212	254
Total	14767	7492	3013	4262

Abbreviation: SLE, systemic lupus erythematosus.

population of EA cases and healthy controls. Two independent effects account for this association, located in the intronic region approximately 31 kb away from the 3′ end of *MYH9*. Unlike previous studies in AAs, this signal is not explained by variants within the neighboring *APOL1* gene. The two *APOL1* coding variants accounting for association between *MYH9* and renal disease in AAs were monomorphic in the EA sample as they are known to be present in very low frequencies ($<0.5\%$) in general EA populations. Moreover, results of our analyses of SLE patients with SLE-related renal dysfunction and thrombocytopenia suggest a broader involvement of *MYH9* in lupus complications. Finally, this report identified the first evidence of suggestive association between *APOL1* and LN with a nominally significant $P<0.05$ in AAs. Our results highlight the complex behavior of a single gene across multiple disorders and racial groups, suggesting the need for additional genetic and combined gene-environment studies.

MATERIALS AND METHODS

Study populations and SNP genotyping

Independent study participants were obtained through 19 national and international collaborators as part of the Large Lupus Association Study 2 (LLAS2). Their respective Institutional Review Boards approved all recruitment studies. Only subjects who signed informed consent forms were included in the study. All SLE patients fulfilled the revised 1997 American College of Rheumatology for classification of SLE[5] and satisfied the renal criterion of either (1) persistent proteinuria >0.5 g per day (24 h) or persistent $>3+$ if quantification was not performed or (2) presence of urinary cellular casts.[5] The LLAS2 study included 8922 SLE cases, 3212 of which fulfilled the renal ACR SLE criterion (Supplementary Table 1) and 4505 were classified as SLE without renal complication, thus comprising the sample analyzed in this report and referred to as renal cases and non-renal SLE cases, respectively. Renal failure documentation based on medical record information of dialysis and/or kidney transplantation identified a subset of 115 patients with severe LN. The control population consisted of 8077 unrelated, healthy, population-based controls with no blood relatives with SLE, bringing the total subjects studied herein to 15 794.

A total of 78 *MYH9* SNPs, including eight previously associated with ESRD,[13] were genotyped in 7717 SLE cases and 8077 healthy controls from six different ethnic groups: EA, AA, AS, HI, Amerindian and Gullah (Supplementary Table 1). The Gullah are a population of AAs residing in the coastal regions of South Carolina and Georgia who exhibit both unique African ancestral origins and lesser European admixture.[27] Note, that all SNPs in the NCBI 36 database, within the *MYH9* region were submitted for inclusion in our custom genotyping assay. The 78 SNPs presented here were those that met Illumina QC standards. All SNPs were in moderate LD (r^2 or $D'<0.80$) with one another. Data were generated using custom designed Illumina iSelect Infinium II genotyping arrays on the BeadStation iScan (Illumina, San Diego, CA, USA) at the Oklahoma Medical Research Foundation (OMRF). In addition, two *APOL1* H-ESRD and FSGS risk variants (rs73885319 (G1) and rs71785313 (G2))[17] were genotyped using custom TaqMan SNP genotyping Assays (Supplementary Method 1) in a subset of 407 AA (205 LN cases and 202 healthy controls) and 579 EA subjects (205 LN cases and 374 healthy controls) from the above cohort for which additional DNA was available.

Quality control

To perform global ancestry estimation, a panel of 347 genomic ancestry informative markers (Supplementary Table 2) was genotyped[28,29] to evaluate the population ancestry and any possible hidden population substructure. The SNPs available in the *MYH9* region are 650 kb away from the nearest ancestry informative markers, and we were therefore unable to accurately estimate the local ancestry in this region.

SNPs included in the analysis had a call rate $>90\%$, $P>0.001$ for Hardy–Weinberg proportion in controls, and minor allele frequencies >0.001. Samples with low call rate ($<90\%$), sample heterozygosity outliers (>5 standard deviations from the mean), extreme population outliers (based on global ancestry estimation and principal component analysis), sample duplicates (proportion of alleles shared identity-by-descent >0.4) and gender discrepancy between reported gender and genetic data were excluded from analysis (Supplementary Method 2 and Supplementary Table 3). After quality control, the final data set comprised 77 *MYH9* SNPs, 2 *APOL1* SNPs, 262 ancestry informative markers, 3013 LN cases, 4262 non-LN SLE cases and 7492 healthy controls (Table 3).

Ancestry estimation

We performed global ancestry estimation for every individual in our study using ADMIXMAP.[30–32] This software adopts a combination of classical and

Bayesian frameworks, and calculated ancestry information through a Markov Chain Monte Carlo (MCMC) simulation using 262 ancestry informative markers and the allele frequencies obtained from the HapMap release 27. Global ancestry estimates were computed for AS, European, Amerindian and West African ancestries.

Imputation method

Imputation was performed over a 105 kb interval flanking the *MYH9* gene on chromosome 22 from 35 Mb to 35.15 Mb using IMPUTE2.[33-35] A collection of 77 SNPs was used as the source of observed genotypes and data from the 1000 Genomes Project and the Phase III HapMap release 2 were used as the reference panels. IMPUTE2 computes posterior probabilities for the three possible genotypes (that is, AA, AB and BB) and then converts posterior probabilities to the most likely genotypes with a threshold of 0.9. Imputed SNPs with low imputation accuracy (information measure <0.5 and $<90\%$ average certainty of the most probable genotypes) were removed from the analysis. We chose to call genotypes so as to be able to construct haplotypes and calculate LD. We did, however, verify that this was indeed a conservative approach by also analyzing SNP 'dose' using SNPTEST[36-38] (Supplementary Table 10). After imputation and quality control evaluation, as described above, each data set comprised a minimum of 89 SNPs for each of the populations (the numbers varied based on LD structure) and is shown in Supplementary Table 4.

Association analysis

To investigate the genotype–phenotype relationship of *MYH9* and *APOL1* polymorphisms in different racial groups, logistic regression including adjustment for gender and global ancestry (quantified in terms of European, African and AS ancestry) was performed to test for association for *MYH9* and *APOL1* SNPs assuming additive, dominant and recessive modes of inheritance using PLINK.[17-19,39,40] We performed analyses in two ways: (1) using healthy population-based participants as controls and (2) using SLE, non-renal cases as controls. However, because the results were not significantly different, we concentrate those from the former, much larger, data set in the main text. All reported Wald χ^2 *P*-values, 95% confidence intervals and ORs were calculated from the logistic regression model. We controlled for experiment-wide type I error by establishing Bonferroni correction thresholds for significance of 2.03×10^{-3} for *MYH9* and 3×10^{-2} for *APOL1*, based on the maximum average number of tests across all populations and weighted for non-independence (that is, $D' > 0.80$). Pair-wise LD measures for the *MYH9* and *APOL1* SNPs were assessed by the D′ values using Haploview 4.2.[41] Finally, we conducted a conditional association analysis using PLINK,[17-19,39,40] adjusting for gender and global ancestry to determine whether the effects seen in EA were independent.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank all study participants, SLE and controls in this study as well as all the staff who assisted in their recruitment. We gratefully acknowledge the following individuals for their generous contribution in genotyping samples: Dr Peter K Gregersen, Drs Sandra D'Alfonso (Italy), Rafaella Scorza (Italy), Peter Junker and Helle Laustrup (Denmark), Marc Bijl (Holland), Emoke Endreffy (Hungary), Carlos Vasconcelos and Berta Martins da Silva (Portugal), Ana Suarez and Carmen Gutierrez (Spain), Iñigo Rúa-Figueroa (Spain) and Dr Cintia Garcilazo (Argentina). For the AADEA collaboration: Norberto Ortego-Centeno (Spain), Juan Jimenez-Alonso (Spain), Enrique de Ramon (Spain) and Julio Sanchez-Roman (Spain). For the GENLES collaboration: Dr Mario Cardiel (Mexico), Dr Ignacio García de la Torre (Mexico), Marco Maradiaga (Mexico), José F Moctezuma (Mexico), Dr Eduardo Acevedo (Peru), Cecilia Castel and Mabel Busajm (Argentina), Jorge Musuruana (Argentina). Other participants from the Argentine Collaborative Group are: Hugo R Scherbarth MD, Pilar C Marino MD, Estela L Motta MD; Susana Gamron MD, Cristina Drenkard MD, Emilia Menso MD; Alberto Allievi MD, Guillermo A Tate MD; Jose L Presas MD; Simon A Palatnik MD, Marcelo Abdala MD, Mariela Bearzotti PhD; Alejandro Alvarellos MD, Francisco Caeiro MD, Ana Bertoli MD; Sergio Paira MD, Susana Roverano MD; Cesar E Graf MD, Estela Bertero PhD; Carolina Guillerón MD, Sebastian Grimaudo PhD, Jorge Manni MD; Luis J Catoggio MD, Enrique R Soriano MD, Carlos D Santos MD; Cristina Prigione MD, Fernando A Ramos MD, Sandra M Navarro MD; Guillermo A Berbotto MD, Marisa Jorfen MD, Elisa J Romero PhD; Mercedes A Garcia MD, Juan C Marcos MD, Ana I Marcos MD; Carlos E Perandones MD, Alicia Eimon MD; Cristina G Battagliotti MD. We also would like to knowledge Mary C Comeau MA; Miranda C Marion MA; Paula S Ramos PhD; Summer Frank MPH and Mai Li Zhu MS for their assistance in genotyping, quality control analyses and clinical data management, and everyone at the Lupus Family Registry and Repository (LFRR) for data collection and maintenance. The work has been funded principally by the US National Institutes of Health grants R01 AI063274 and R01 AR056360 (PMG); R01 AR043274 (KLM); N01 AR62277, R37 24717, R01 AR042460, P01 AI083194, P20 RR020143, R01 DE018209 (JBH); P01 AR49084 (RPK and EEB); R01 AR33062 (RPK); P30 AR055385 (EEB); K08 AI083790, LRP AI071651, UL1 RR024999 (TBN); R01CA141700, RC1 AR058621 (MEAR); R01AR051545-01A2, ULI RR025014-02 (AMS); P30 AR053483, N01 AI50026 (JAJ and JMG); P20 RR015577 (JAJ); R21 AI070304 (SAB); P30 RR031152, U19 AI082714, P30 AR053483, RC1 AR05884 (JAJ and JMG); R01 AR43814 (BPT); P60 AR053308, M01 RR-00079 (LAC); R01 AR043727, UL1 RR025005 (MAP); K24 AR002138, P60-2 AR30692, P01 AR49084, UL1RR025741 (RRG); UL1 RR029882, P60 AR049459 (GSG and DLK). A portion of this study was supported by a grant of the Korea Healthcare Technology R&D Project, Ministry for Health and Welfare, Republic of Korea (A080588; SCB). Additional support was granted from the Alliance for Lupus Research (KLM); Merit Award from the US Department of Veterans Affairs (JBH and GSG); the Swedish Research Council for Medicine, Gustaf Vth-80th Jubilee Fund and Swedish Association Against Rheumatism, Instituto de Salud Carlos III, Oklahoma Center for Advancement of Science and Technology (OCAST) HR09-106 (MEAR); the European Science Foundation funds the BIOLUPUS network (MEAR coordinator); Federico Wilhelm Agricola Foundation Research grant (BPE); The Barrett Scholarship Fund OMRF (CJL); Lupus Research Institute (TBN, BPT); The Alliance for Lupus Research (TBN, LAC, MEAR and COJ); the Arthritis National Research Foundation Eng Tan Scholar Award (TBN); Arthritis Foundation (PMG and AMS); the Lupus Foundation of Minnesota (PMG and KLM); the Wellcome Trust (TJV); Arthritis Research UK (TJV); Kirkland Scholar Award (LAC, JAJ) and Wake Forest University Health Sciences Center for Public Health Genomics (CDL). The work reported on in this publication has been in part financially supported by the ESF, in the framework of the Research Networking Programme European Science Foundation-The Identification of Novel Genes and Biomarkers for Systemic Lupus Erythematosus (BIOLUPUS) 07-RNP-083.

REFERENCES

1 Hart HH, Grigor RR, Caughey DE. Ethnic difference in the prevalence of systemic lupus erythematosus. *Ann Rheum Dis* 1983; **42**: 529-532.
2 Bae SC, Fraser P, Liang MH. The epidemiology of systemic lupus erythematosus in populations of African ancestry: a critical review of the 'prevalence gradient hypothesis'. *Arthritis Rheum* 1998; **41**: 2091-2099.
3 Serdula MK, Rhoads GG. Frequency of systemic lupus erythematosus in different ethnic groups in Hawaii. *Arthritis Rheum* 1979; **22**: 328-333.
4 Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, Rothfield NF *et al.* The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1982; **25**: 1271-1277.
5 Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997; **40**: 1725.
6 Wallace DJ, Hahn B, Dubois EL. *Dubois' Lupus Erythematosus* 6th edn, Lippincott Williams & Wilkins: Philadelphia, 2002, xiii, 1348 pp.
7 Cameron JS. Lupus nephritis. *J Am Soc Nephrol* 1999; **10**: 413-424.
8 Cameron JS. Lupus nephritis: an historical perspective 1968-1998. *J Nephrol* 1999; **12**(Suppl 2): S29-S41.
9 Bastian HM, Roseman JM, McGwin Jr G, Alarcon GS, Friedman AW, Fessler BJ *et al.* Systemic lupus erythematosus in three ethnic groups. XII. Risk factors for lupus nephritis after diagnosis. *Lupus* 2002; **11**: 152-160.
10 Mok CC, Lau CS. Lupus in Hong Kong Chinese. *Lupus* 2003; **12**: 717-722.
11 Wong SN, Tse KC, Lee TL, Lee KW, Chim S, Lee KP *et al.* Lupus nephritis in Chinese children–a territory-wide cohort study in Hong Kong. *Pediatr Nephrol* 2006; **21**: 1104-1112.
12 Freedman BI, Hicks PJ, Bostrom MA, Cunningham ME, Liu Y, Divers J *et al.* Polymorphisms in the non-muscle myosin heavy chain 9 gene (MYH9) are strongly associated with end-stage renal disease historically attributed to hypertension in African Americans. *Kidney Int* 2009; **75**: 736-745.
13 Kao WH, Klag MJ, Meoni LA, Reich D, Berthier-Schaad Y, Li M *et al.* MYH9 is associated with nondiabetic end-stage renal disease in African Americans. *Nat Genet* 2008; **40**: 1185-1192.

14 Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW *et al.* MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet* 2008; **40**: 1175 - 1184.
15 Behar DM, Rosset S, Tzur S, Selig S, Yudkovsky G, Bercovici S *et al.* African ancestry allelic variation at the MYH9 gene contributes to increased susceptibility to non-diabetic end-stage kidney disease in Hispanic Americans. *Hum Mol Genet* 2010; **19**: 1816 - 1827.
16 Kunishima S, Saito H. Advances in the understanding of MYH9 disorders. *Curr Opin Hematol* 2010; **17**: 405 - 410.
17 Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, Freedman BI *et al.* Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 2010; **329**: 841 - 845.
18 Genovese G, Tonna SJ, Knob AU, Appel GB, Katz A, Bernhardy AJ *et al.* A risk allele for focal segmental glomerulosclerosis in African Americans is located within a region containing APOL1 and MYH9. *Kidney Int* 2010; **78**: 698 - 704.
19 Zenker M, Mertens PR. Arrest of the true culprit and acquittal of the innocent? Genetic revelations charge APOL1 variants with kidney disease susceptibility. *Int Urol Nephrol* 2010; **42**: 1131 - 1134.
20 Freedman BI, Edberg JC, Comeau ME, Murea M, Bowden DW, Divers J *et al.* The non-muscle myosin heavy chain 9 gene (MYH9) is not associated with lupus nephritis in African Americans. *Am J Nephrol* 2010; **32**: 66 - 72.
21 Saban Elitok UG, Markus B, Mato N, Wolfgang S, Ralph K, Friedrich CL. MYH9 mutation and lupus erythematosus. *NDT Plus* 2010; **3**: 128 - 131.
22 Reville K, Crean JK, Vivers S, Dransfield I, Godson C. Lipoxin A4 redistributes myosin IIA and Cdc42 in macrophages: implications for phagocytosis of apoptotic leukocytes. *J Immunol* 2006; **176**: 1878 - 1888.
23 Marigo V, Nigro A, Pecci A, Montanaro D, Di Stazio M, Balduini CL *et al.* Correlation between the clinical phenotype of MYH9-related disease and tissue distribution of class II nonmuscle myosin heavy chains. *Genomics* 2004; **83**: 1125 - 1133.
24 Mok CC. Biomarkers for lupus nephritis: a critical appraisal. *J Biomed Biotechnol* 2010; **2010**: 638413.
25 Murea M, Freedman BI. Essential hypertension and risk of nephropathy: a reappraisal. *Curr Opin Nephrol Hypertens* 2010; **19**: 235 - 241.
26 Nelson GW, Freedman BI, Bowden DW, Langefeld CD, An P, Hicks PJ *et al.* Dense mapping of MYH9 localizes the strongest kidney disease associations to the region of introns 13 to 15. *Hum Mol Genet* 2010; **19**: 1805 - 1815.
27 Kamen DL, Barron M, Parker TM, Shaftman SR, Bruner GR, Aberle T *et al.* Autoantibody prevalence and lupus characteristics in a unique African American population. *Arthritis Rheum* 2008; **58**: 1237 - 1247.
28 Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A *et al.* A high-density admixture map for disease gene discovery in african americans. *Am J Hum Genet* 2004; **74**: 1001 - 1013.
29 Halder I, Shriver M, Thomas M, Fernandez JR, Frudakis T. A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: utility and applications. *Hum Mutat* 2008; **29**: 648 - 658.
30 McKeigue PM, Carpenter JR, Parra EJ, Shriver MD. Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: application to African-American populations. *Ann Hum Genet* 2000; **64**(Part 2): 171 - 186.
31 Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, Clayton DG *et al.* Control of confounding of genetic associations in stratified populations. *Am J Hum Genet* 2003; **72**: 1492 - 1504.
32 Hoggart CJ, Shriver MD, Kittles RA, Clayton DG, McKeigue PM. Design and analysis of admixture mapping studies. *Am J Hum Genet* 2004; **74**: 965 - 978.
33 Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009; **5**: e1000529.
34 Via M, Gignoux C, Burchard EG. The 1000 Genomes Project: new opportunities for research and social challenges. *Genome Med* 2010; **2**: 3.
35 Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**: 851 - 861.
36 Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet* 2010; **11**: 499 - 511.
37 Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; **39**: 906 - 913.
38 Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature* 2007; **447**: 661 - 678.
39 Freedman BI, Kopp JB, Langefeld CD, Genovese G, Friedman DJ, Nelson GW *et al.* The apolipoprotein L1 (APOL1) gene and nondiabetic nephropathy in African Americans. *J Am Soc Nephrol* 2010; **21**: 1422 - 1426.
40 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559 - 575.
41 Barrett JC. Haploview: visualization and analysis of SNP genotype data. *CSH Protoc* 2009; **2009**: pdb ip71.

Supplementary Information accompanies the paper on Genes and Immunity website (http://www.nature.com/gene)